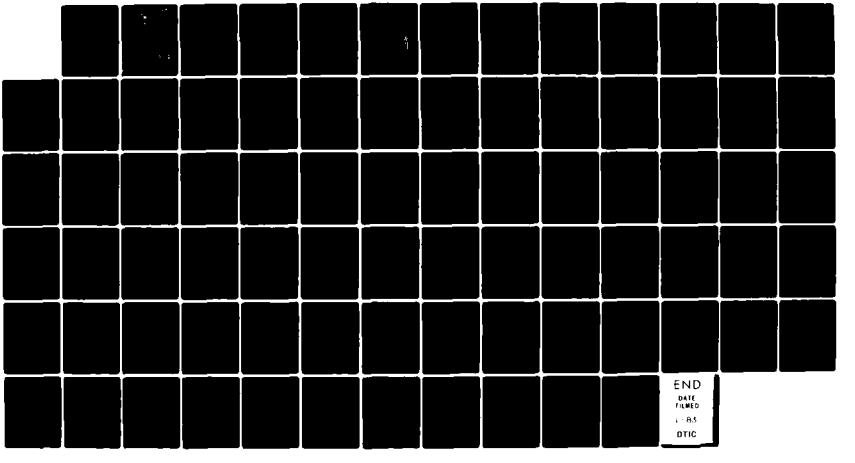
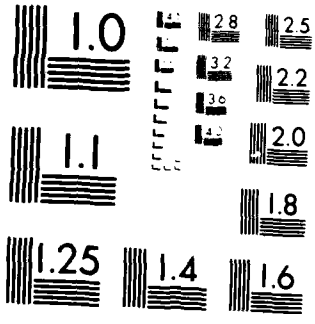
AD-A123 005 PREDICTIVE MEASURES FOR THE ACHIEVEMENT OF TRAINING SUCCESS IN AIR FORCE..(U) AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOOL OF SYST.. 1/1

UNCLASSIFIED T M NEWSTAD ET AL. SEP 82 AFIT-LSSR-37-82 F/G 5/9 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

AD A123005

DTIC
ELECTE
S JAN 6 1983 D
D

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY (ATC)

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

83 1 5 061

DTIC FILE COPY

Accession For
NTIS CRA&I
DTIC TAB
Unannounced
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A

PREDICTIVE MEASURES FOR THE ACHIEVEMENT OF TRAINING SUCCESS IN AIR FORCE TECHNICAL TRAINING

Theodore M. Newstad, Captain, USAF
James A. Schuster, Captain, USAF

LSSR 37-82

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

The contents of the document are technically accurate, and no sensitive items, detrimental ideas, or deleterious information are contained therein. Furthermore, the views expressed in the document are those of the author(s) and do not necessarily reflect the views of the School of Systems and Logistics, the Air University, the Air Training Command, the United States Air Force, or the Department of Defense.

AFIT Control Number LSSR 37-82

AFIT RESEARCH ASSESSMENT

The purpose of this questionnaire is to determine the potential for current and future applications of AFIT thesis research. Please return completed questionnaires to: AFIT/LSH, Wright-Patterson AFB, Ohio 45433.

1. Did this research contribute to a current Air Force project?

   a. Yes b. No

2. Do you believe this research topic is significant enough that it would have been researched (or contracted) by your organization or another agency if AFIT had not researched it?

   a. Yes b. No

3. The benefits of AFIT research can often be expressed by the equivalent value that your agency received by virtue of AFIT performing the research. Can you estimate what this research would have cost if it had been accomplished under contract or if it had been done in-house in terms of manpower and/or dollars?

   a. Man-years _____ $ _____ (Contract).

   b. Man-years _____ $ _____ (In-house).

4. Often it is not possible to attach equivalent dollar values to research, although the results of the research may, in fact, be important. Whether or not you were able to establish an equivalent value for this research (3 above), what is your estimate of its significance?

   a. Highly Significant b. Significant c. Slightly Significant d. Of No Significance

5. Comments:

_____ _____

Name and Grade Position

_____ _____

Organization Location

FOLD DOWN ON OUTSIDE - SEAL WITH TAPE

AFIT/ LSH
WRIGHT-PATTERSON AFB OH 45433

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE. $300

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL
FIRST CLASS PERMIT NO. 73236 WASHINGTON D.C.

POSTAGE WILL BE PAID BY ADDRESSEE

AFIT/ DAA
Wright-Patterson AFB OH 45433

FOLD IN

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: LSSR 37-82	2. GOVT ACCESSION NO.: AD-A123 005
	3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): PREDICTIVE MEASURES FOR THE ACHIEVEMENT OF TRAINING SUCCESS IN AIR FORCE TECHNICAL TRAINING	5. TYPE OF REPORT & PERIOD COVERED: Master's Thesis
	6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): Theodore M. Newstad, Captain, USAF James A. Schuster, Captain, USAF	8. CONTRACT OR GRANT NUMBER(s):
9. PERFORMING ORGANIZATION NAME AND ADDRESS: School of Systems and Logistics Air Force Institute of Technology, WPAFB OH	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS:
11. CONTROLLING OFFICE NAME AND ADDRESS: Department of Communication and Humanities AFIT/LSH, WPAFB OH 45433	12. REPORT DATE: September 1982
	13. NUMBER OF PAGES: 72
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office):	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

APPROVED FOR PUBLIC RELEASE: IAW AFR 190-17
AIR FORCE INSTITUTE OF TECHNOLOGY (ATC)
WRIGHT-PATTERSON AFB, OH 45433
8 OCT 1982

LYNN E. WOLAVER
Dean for Research and
Professional Development

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Validity of Instructional Measures
Validity of Biographical Measures
Validity of Psychological Tests
Validity of Intelligence Tests
Prediction of Success

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Thesis Chairman: Robert P. Steel, Ph.D., GM-13

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

This research was conducted to investigate the predictive nature of a variety of measures for the achievement of training success at the Air Force Technical Training School, Chanute Air Force Base, Illinois. A sample of 1358 male and female airmen attending twelve technical training courses from 1 July 1981 to 15 February 1982 was analyzed using multiple regression analysis. Final course grade served as the criterion of training success. Predictor variables included four aptitude indices and one intelligence score from the Armed Services Vocational Aptitude Battery (ASVAB), seven biographical items, and four on-site instructional variables. Separate regressions were carried out on the entire sample and for each individual technical training course. The results were discussed in terms of their implications not only for the classification of students to training, but also for the remediation of problem students. Findings showed that test scores and on-site instructional variables significantly added to the prediction of the final course grade while biographical items demonstrated only moderate predictive validity for the criterion.

LSSR 37-82

PREDICTIVE MEASURES FOR THE ACHIEVEMENT OF TRAINING SUCCESS IN AIR FORCE TECHNICAL TRAINING

A Thesis

Presented to the Faculty of the School of Systems and Logistics

of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Systems Management

By

Theodore M. Newstad, BS
Captain, USAF

James A. Schuster, MS
Captain, USAF

September 1982

Approved for public release;
distribution unlimited

This thesis, written by

Captain Theodore M. Newstad

and

Captain James A. Schuster

has been accepted by the undersigned on behalf of the faculty of the School of Systems and Logistics in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN SYSTEMS MANAGEMENT

DATE: 29 September 1982

Robert P. Steel
COMMITTEE CHAIRMAN

READER

TABLE OF CONTENTS

	Page
LIST OF TABLES	vi
Chapter	
1. INTRODUCTION	1
Problem Statement	1
Justification	1
Background	2
Intelligence and Aptitude Test Development	3
Biographical Items as Predictors	6
Biographical Material and Military Test Batteries	10
Instructional Variables	16
Research Objectives	16
Scope and Limitations	17
2. METHOD	18
Population and Sample	18
Criterion Variable	19
Predictor Variables	19
Pre-Training Variables	20
Mechanical Score	20
Administrative Score	20
General Score	20
Electronic Score	20

Chapter Page

Intelligence Score 20

Age 21

Sex 21

First Choice 21

Military Dependent 21

Student Source 21

Education Level 21

Instructional Variables 22

Shift 22

Hours Absent 22

Tutoring 22

Academic Counselings 23

Nonacademic Counselings 23

Procedures 23

Analytic Techniques 24

Pearson Product-Moment Correlation 24

Multiple Regression 25

Stepwise Multiple Regression 25

3. RESULTS 29

Pearson Product-Moment Correlations 29

Stepwise Multiple Regression 29

All Courses Combined 31

Nondestructive Inspection Specialist Course 33

Fabrication and Praachute Specialists Course 34

Chapter Page

Airframe Repair Specialist Course 35

Special Vehicle Maintenance Courses 37

General Purpose Vehicle Mechanic Course . . 37

Fire Protection Specialist Course 38

Aircrew Life Support Specialist Course . . . 40

Evaluation of Hypothesis 1 41

Evaluation of Hypothesis 2 42

Evaluation of Hypothesis 3 42

4. DISCUSSION, CONCLUSIONS, AND RECOMMENDATIONS . . 44

Test Scores 44

Instructional Variables 47

Biographical Items 49

Conclusions and Recommendations 50

APPENDICES 54

A. REGRESSION EQUATIONS 55

B. TECHNICAL TRAINING COURSE SELECTOR SCORES . . . 57

C. TECHNICAL TRAINING SURVEY WORKSHEET 59

BIBLIOGRAPHY 63

LIST OF TABLES

Table		Page
1.	Intercorrelations Among Predictor and Criterion Variables	30
2.	Regression Over All AFSCs with Final Course Grade as the Criterion Variable (N=1358)	32
3.	Regression for AFSC 42732 with Final Course Grade as the Criterion Variable (N=80)	34
4.	Regression for AFSC 42733 with Final Course Grade as the Criterion Variable (N=60)	35
5.	Regression for AFSC 42735 with Final Course Grade as the Criterion Variable (N=296)	36
6.	Regression for AFSCs 47231A, 47231B, 47231C, 47231D with Final Course Grade as the Criterion Variable (N=84)	38
7.	Regression for AFSC 47232 with Final Course Grade as the Criterion Variable (N=56)	39
8.	Regression for AFSC 57130 with Final Course Grade as the Criterion Variable (N=541)	39
9.	Regression for AFSC 12230 with Final Course Grade as the Criterion Variable (N=206)	41

CHAPTER 1

INTRODUCTION

Problem Statement

This study is being conducted to determine the predictive nature of a variety of measures for the achievement of training success at the Air Force Technical Training School, Chanute Air Force Base, Illinois.

Justification

Previous research has been conducted primarily to determine how well an airman's aptitude index, as derived from the Armed Services Vocational Aptitude Battery (ASVAB) given in the high schools, predicts technical school success (final course grade). While a handful of isolated investigations have examined the use of education level, age, sex, and other biographic/demographic measures to predict final course grade, more extensive investigations of these measures can replicate and extend Air Force research examining their feasibility as supplemental predictors to ASVAB test scores. The accuracy of personnel assignment and classification decisions may be thus enhanced leading to improved student performance and lower course attrition rates for enlisted military personnel attending Air Force technical training.

Background

Psychological tests can be valuable instruments for accurately providing information about an individual's capacities and characteristics. For the most part, tests are objective, standardized measures of a sample of human behavior which can be used to assess the differences among individuals or between one individual's reactions on two different occasions (Anastasi, 1968:24).

Originally developed to identify individuals who were mentally retarded or otherwise mentally deficient, intelligence tests have evolved into a widely-used means of resolving classification problems in education and employment. In this regard, psychological tests have been found to be especially useful in personnel selection and placement, in vocational and educational counseling, and in the diagnosis of academic failures (Anastasi, 1968:3).

Psychological tests have proven effective in the selection and classification of industrial personnel. When scores from such tests are validated, they can serve as predictors of job proficiency, labor turnover, or job training. When the prediction of job training is of primary interest, the most common indices of success are training course grades and instructor ratings (Ghiselli and Brown, 1955:187,219).

According to Anastasi (1968), the usefulness of a psychological test as a predictive measure depends

primarily upon "the degree to which it serves as an indicator of a relatively broad and significant area of behavior [Anastasi, 1968:22]." Thus, to be effective, tests must accurately indicate an individual's "mastery" of a particular behavior from sample information gathered via the testing instrument.

Intelligence and Aptitude Test Development

In 1904 Binet, along with his coworkers, provided the initial impetus for the development of intelligence tests. One important outcome of their research efforts was the Binet-Simon scales used to measure mental age.

The Binet-Simon tests were revised in 1916 by L. M. Terman of Stanford University. These tests used a ratio of mental age and chronological age (IQ) as an index of intelligence (Anastasi, 1968:11).

With the onset of World War II, psychologists recognized the need to accomplish large group testing to classify over a million and a half recruits with respect to IQ. After reviewing a large amount of published and unpublished material on testing, Army psychologists developed the Army Alpha and Army Beta. These tests were found to be suitable for administration to large groups and served as models for group intelligence testing in general (Anastasi, 1968:12).

In the twenties, the name "intelligence test" gradually came to be recognized as a misnomer, due to the limited information it yielded. Many tests were instead termed "scholastic aptitude tests" because they were designed to measure a more specific combination of abilities related to academic performance.

The military continued with separate lines of development for intelligence tests and aptitude tests with the respective development of the Armed Forces Qualifying Test, used for initial selection decisions, and the General Classification Test, used for classification decisions. These efforts converged in the development of multiple aptitude batteries designed to provide a measure of an individual's standing on a number of traits (Anastasi, 1968:14).

Today, the Air Force military testing system exists primarily to meet the need for timely personnel selection and placement decisions by "providing selection, classification and assignment, job knowledge, and military knowledge tests [U.S. Department of the Air Force, 1978:p.1-1]." This system relies heavily, as do the other services' testing systems, on a single enlisted selection and classification battery known as the Armed Services Vocational Aptitude Battery (ASVAB). Administered in high schools and vocational trade schools, ASVAB yields four aptitude scores--mechanical, administrative, general, and electronic--

and an intelligence score. This score is the IQ equivalent of the now obsolete Armed Forces Qualification Test (AFQT) which was previously used for selection decisions.

Forms of the ASVAB currently in use (8, 9, and 10) became operational in October, 1980, and are comprised of ten subtests: general science (GS), arithmetic reasoning (AR), word knowledge (WK), paragraph comprehension (PC), numerical operations (NO), coding speed (CS), autoshop information (AS), mathematics knowledge (MK), mechanical comprehension (MC), and electronics information (EI) (Ree, Mullins, Mathews, and Massey, 1982:29). These categories were chosen to enable each branch of the service to merge these subtests into four aptitude indices similar to those derived from the test batteries each service previously used. Using these aptitude indices, the services can assign recruits into one of the four broad aptitude clusters.

The intelligence score is currently derived from the AR, WK, PC, and NO subtests and is designed to yield a measure of academic ability which aids in initial selection, rather than assignment, decisions (Ree et al., 1982:7; Valentine, 1977:10).

Since the introduction of Form 1, ASVAB has continued to be validated, revised, and developed to keep pace with the latest psychometric advances and to meet enlisted selection and classification requirements of all the Armed

Forces (Vitola and Alley, 1968; Vitola, Mullins, and Croll, 1973; Weeks, Mullins, and Vitola, 1975; Jensen, Massey and Valentine, 1976; Mullins, Ehrles, and Ree, 1981). Nevertheless, these "custom-built" tests remain imperfect, being in themselves incapable of providing a totally comprehensive measure of potential for training success. According to AFR 35-8, regarding such personnel testing, one ought never to "use tests to the exclusion of other information [U.S. Department of the Air Force, 1978: p.1-1]." In other words, the predictive value of other types of measures should not be overlooked in the classification and assignment of airmen to technical training.

To illustrate the value of alternative measures in the prediction of tenure, turnover, and training success, we will first present a sample of investigations validating biographical data in these kinds of prediction situations and then review a representative sampling of studies validating military test batteries in conjunction with biographical items in personnel selection and classification decisions. The review will close with a discussion of "instructional" variables and their use for the prediction of training success.

Biographical Items as Predictors

Ehrle (1964) designed a biographical questionnaire for predicting job success for a sample of 200 job

applicants. Using eighty-six items of personal data to form one classification key, and twenty selected items, such as age, education, marital and employment status, he found that, even in the cross-validation groups, there was an increase in predictive efficiency of 10 to almost 19 percent over chance when using biographical keys to predict success or failure for the applicants. He further noted that the twenty-variable key predicted vocational success better than the eighty-six-variable key with 67.5 percent versus 60.5 percent of the applicants correctly classified in the primary (N=200) sample and 68.8 percent versus 65.0 percent correctly classified in the cross-validation sample (N=80). This result was attributed to the shorter key being more closely related to the dichotomous criterion for state vocational rehabilitation services than was the longer key (Ehrle, 1964:174).

Cascio (1976) also cross-validated biographical scoring keys for a set of biographical data. The objective of his investigation was to predict tenure of 160 minority and non-minority female clerical personnel in a medium-sized insurance company from ten items which previous research had shown to be valid predictors of tenure. The analysis resulted in a measure of the same ten items surviving for both subgroups; age, marital status, education, children's ages, tenure of previous job, previous salary,

friend or relative within the company, location of residence, home ownership, and length of time at present address.

Weighted combinations of these ten application-blank items yielded significant cross-validated rs of .58 (minorities) and .56 (non-minorities). Cascio also discovered that, even after satisfying the legal restrictions applying to the use of such information, accurate forecasts of turnover could be made by utilizing custom-tailored keys consisting of biographical data (Cascio, 1976:578-579).

Stone and Athelson (1969) researched the prediction of women's "occupational tenure" for 198 occupational therapists and 255 physical therapists. Using eight demographic predictors and selected scales of the women's Strong Vocational Interest Blank (SVIB), the authors employed a double cross-validation design and two types of prediction equations. From the eight demographic items, marital status correlated significantly with tenure for the physical therapists (-.47 and -40), although it was not predictive for the occupational therapists. Number of children, age of children, personal income, and spouse's income were the best predictors of tenure when used uniquely with both occupational and physical therapist samples. None of the SVIB scales proved to be stable predictors in this study (Stone and Athelson, 1969:408,410).

McClelland and Rhodes (1969) compared biographical data with the Minnesota Multi-Phasic Personality Inventory (MMPI) for predicting the job success of 111 hospital aides and 100 orderlies. Age, marital status, number of dependents, and years of education were among the thirteen biographical items surveyed.

Of the variables investigated, the marital status classification, divorced, yielded the highest correlation (.50) with the criterion of job success of all variables studied. Furthermore, this same variable correlated with five additional criterion measures, "making it the most useful biographical predictor and among the most useful of all predictors [McClelland and Rhodes, 1969:54]."

Although education was not significantly correlated with any individual job performance measure, the researchers noted that education did appear to be one of the most effective variables in predicting the composite criterion of job success, particularly for orderlies. Overall, comparing biographical measures to the MMPI, the authors concluded that "biographical variables were more important than MMPI scores for predicting composite job performance [McClelland and Rhodes, 1969:54]."

Asher (1972) reviewed the validity of biographical items when used in combination as predictors of skilled and unskilled work behavior. He compared the average size of biographical validity coefficients to those typically

obtained with other selection procedures (e.g., aptitude tests). The items appeared on scorable application blanks for eleven selected studies between 1960 and 1970. Asher found that, of the thirty-one cross-validated coefficients available, 35 percent were .60 or higher, 55 percent were .50 or higher, 74 percent were .40 or higher, and 97 percent were .30 or higher. Summarizing his results, Asher concluded that

> . . . in comparison with other predictors [such] as intelligence, aptitude, interest, and personality [measures], biographical items had vastly superior validity [Asher, 1972:266].

Finally, in a review of the pertinent literature available on the predictability of employee turnover, Schuh (1967) provides evidence corroborating Asher's conclusions concerning the utility of biographical items for predicting work outcomes. In nineteen of the twenty-one validation studies reviewed, one or more of the items predicted a criterion of labor turnover supporting, in the author's judgment, the rationale for inclusion of personal history items in conjunction with other predictors (Schuh, 1967:141).

Biographical Material and Military Test Batteries

Lecznar (1964) used a multiple regression technique to analyze and to assess percentage of contribution of education alone, aptitude tests alone, and then

combinations of aptitude measures, age, and years of education in the prediction of final school grade. He attempted to confirm findings concerning the predictive value of education for Air Force technical training.

In the first of a series of analyses, education, used as a continuous variable of level attained, added nothing to the predictiveness of the administrative, electronics, general, and mechanical ASVAB scores for the prediction of the criterion. In the prediction equation, age and education added only .03 to the squared correlation.

In the second analysis, the largest single increase in the squared correlation was only .05 for years of education in a regression equation in combination with aptitude scores as compared to using a regression equation based only on aptitude scores. The average gain across schools was only .15 (Lecznar, 1964:9-10).

The third analysis revealed minimal value in utilizing even more complex combinations, confirming prior findings: (1) years of education alone as a continuous variable for classification and assignment to training is not warranted, and (2) the educational variable was of negligible importance when added to aptitude data to predict final course grade.

Since education was not treated as a dichotomous variable (graduation versus nongraduation from high school), Lecznar cautioned against generalization of his findings.

Although he focused on education as a continuous variable, Lecznar suggested that dichotomous treatment might increase the efficiency of this variable, as many studies which he reviewed indicated, for the prediction of the success criterion (Lecznar, 1964:10).

Goffman (1975) investigated the use of twenty selected biographical items and five aptitude test scores (similar to ASVAB) for the prediction of paramedic training success for 1448 students graduating from the Naval Hospital Corps School between November, 1966, and August, 1967. Goffman used multiple regression techniques with subtest scores on verbal, arithmetic, clerical, mechanical, and general intelligence from the Navy Basic Test Battery to develop a prediction equation for the training criterion. Training success was expressed as a score from 0 to 9, with 0 representing academic or administrative discharge and 1 through 9 representing graduates "depending on the interval of the gradepoint scale scored [Goffman, 1975:799]."

The most valid predictors of success were reported using Pearson product-moment correlations as follows: general intelligence (r=.56), the arithmetic subtest (r=.43), highest school grade completed (r=.42), age at entry into service (r=.33), age at entry into Corps school training (r=.32), and attitude variables (r=-.23).

With a predictive validity of R=.65 for the total equation, Goffman concluded that, in addition to test

scores, "demographic, personal history, and attitude variables act as unique contributors to account for academic performance criterion variance [Goffman, 1975:802]."

The three-fold objective of a study by Valentine (1977) was to investigate: (1) the validity of ASVAB (version 3) in conjunction with educational data for success in Air Force Technical Training, (2) the unique contribution of both educational and test data in predicting Air Force Technical Training success, and (3) the homogeneity of prediction equations for subgroups defined by race and sex (Valentine, 1977:5).

With data from a sample of 178,170 Air Force non-prior service accessions between September, 1973, and October, 1975, he separately analyzed success in forty-three technical courses using the general intelligence score, the four aptitude scales (mechanical, administrative, general and electronic), forty-one dichotomous variables indicating completion or noncompletion of specific high school courses a dichotomous training variable (graduation or failure), sex, and race.

Validities for final course grade were obtained for the total sample from each course, and for subsamples within each course grouped by race and sex. Separate regression analyses tested the contribution of educational data to test data and then vice versa in prediction of the criterion. Assessment of homogeneity of race and sex

was established for prediction models based on educational data alone, then on test data only, and finally on a combination of the two.

Higher zero-order correlations were obtained for test data than for educational data. Furthermore, validities for the AFQT were as high, and in some cases higher, than those of the four aptitude composites. Predictions based on educational data were more sensitive to race bias than those based on test data, and "race and sex unique predictions based on test and/or educational data are not homogeneous [Valentine, 1977:5]." This last finding suggests that subgroup equations differed enough to warrant a separate equation for race groups and sex groups to achieve added accuracy to the prediction.

One of the objectives of an extensive research effort by Verna and Mifflin (1977) was to determine and recommend to the Deputy Chief of Staff for Manpower at Marine Corps Headquarters improved criteria on which to base assignment of recruits. The eighty-four corps schools studied ranged from food services to avionics. Stepwise multiple regression analysis was used (N=24,380) to determine the predictive validity of the Army Classification Battery (ACB-61), level of education, race, pay, school start dates, school finish dates, and other relevant demographics for the criterion of final course grade.

The best test predictor of training success was the general intelligence score which was derived from the ACB (Army Classification Battery). On the average, it explained greater than 30 percent of the variance in the criterion.

The authors further noted that "knowledge of the recruit's civilian education level (specifically whether he graduated from high school) was often the next best predictor [Verna and Mifflin, 1977:11]." Typically, it additionally explained over 3 percent of the criterion variance. Other aptitude subtests explained about 1 percent of the criterion variance.

From these findings, and from the fact that failure rates were at least three times higher for non-graduates, the authors made two major recommendations having implications for the use of the AFQT mental category scores and educational data:

1. An ASVAB analog of the General Classification Test score should be introduced into the assignment process;
2. All school entrance requirements [should] be made lower for high school graduates than for non-graduates by granting to graduates 10 credit points on all scores used to determine school eligibility [Verna and Mifflin, 1977:xiii].

In reviewing the literature, it is to be noted that studies on civilian employees historically used biographical items to predict tenure and turnover. However, military investigators have not been limited to this apparent

convention and have investigated the training success criterion as well.

Instructional Variables

Although not specifically identified in the literature as common predictors of training success, variables that indicate that a student is experiencing difficulty while training is in progress may prove useful for the prediction of final course grade. If, for instance, certain students in training receive more academic or non-academic counseling than others, or if students are forced to seek additional instruction while in training, such measures may show a negative relationship with the success criterion. Intuitively, one would identify such measures as having predictive strength, and for this reason they have been included in the present investigation. There is some precedent in the professional literature for anticipating relationships between absence frequency, an on-site instructional variable, and measures of performance such as training success (Gussett, 1976; Caffrey and Klugh, 1971).

Research Objectives

In the interest of improving the quality of personnel training decisions, this study is being accomplished to assist the Chief of Evaluation at the Air Force Technical Training School, Chanute AFB, Illinois, in the

investigation of the following hypotheses suggested by the review of the literature:

Hypothesis 1. Aptitude and IQ tests will predict final course grade.

Hypothesis 2. Biographical items will significantly add to the prediction of final course grade.

Hypothesis 3. Instructional variables will add significantly to the prediction of final course grade.

Scope and Limitations

Due to limitations on available data, only seven Air Force technical training schools were included in this investigation for the prediction of training success. Also, due to low course attrition rates, we deemed the pass/fail dichotomy to be a poor criterion for success, and therefore restricted our main analysis to final course grade as a criterion of training success.

Finally, our sample did not include trainees who "washed out" of the training programs prior to the award of final grades. Therefore, the nature of our sample may not allow for generalizability to all individuals entering these technical training schools.

CHAPTER 2

METHOD

This chapter reviews the research methods used, including population and sample, the variables measured, measures, procedures, design, and assumptions underlying use of the independent variables employed to predict final course grade in the Air Force Technical Training School.

Population and Sample

The population consists of all students who have, are, or will be taking any of the following twelve USAF Technical Training Courses at Chanute AFB: 42732, 42733, 42734, 42735, 42730, 47231 (A, B, C, or D), 47232, 57130, 12230. The numbers correspond to Air Force Specialty Codes (AFSCs) (see Appendix B). These courses train airmen as specialists in career fields ranging from maintenance of aircraft systems to maintenance of general and special purpose vehicles.

The sample consisted of 1358 trainees entering one of the above Air Force Specialty Codes (AFSCs) between 1 July 81 and 15 February 82. The mean age for the sample was twenty years ($\bar{x}$ = 20.00). The mean education level for the sample was slightly beyond high school graduate ($\bar{x}$ = 12.21 years of education).

Criterion Variable

Final course grade served as the criterion of training success. This measure denotes the grade received at the end of each student's respective Technical Training Course. Data on students who did not have a final course grade were deleted from the analysis. The remaining students (N = 1358) had a mean course grade of 85.89 with a standard deviation of 8.04.

Predictor Variables

The predictor variables have been divided into Pre-Training and Instructional variables. The Pre-Training variables were constructed from ASVAB (Forms 8, 9, and 10) test scores and biographical material available prior to student training. The Instructional variables were formed from information gathered throughout the process of course instruction and training.

The ASVAB consists of ten aptitude subtests that are merged by the respective services into four aptitude indices and an intelligence score described below. The Air Force Human Resources Laboratory calculated the reliability of these indices using the Wherry-Gaylord method of reliability estimation (Ree, 1982; Wherry and Gaylord, 1943).

Pre-Training Variables

Mechanical Score. This score is obtained from the ASVAB subtests and is used by the Air Force to assign airmen to technical training. The minimum mechanical score for assignment of airmen to training courses in this study is given in Appendix B. The reliability of this measure was calculated as .88.

Administrative Score. This score is derived from the ASVAB subtests and is also used by the Air Force to assign airmen to technical training (see Appendix B). The reliability of this measure was calculated as .70.

General Score. This score is obtained from the ASVAB and is used by the Air Force to assign airmen as indicated in Appendix B. The reliability of this measure was computed as .81.

Electronic Score. This score, also obtained from the ASVAB, is used by the Air Force to assign airmen to technical training (see Appendix B). The reliability of this measure was computed to be .84.

Intelligence Score. This score is constructed from four ASVAB subtests (arithmetic reasoning, word knowledge, paragraph comprehension, and numerical operations) and is commonly used in initial airmen selection rather

than in assignment decisions. This study will examine its usefulness for assignment decisions as well.

Age. The individual's age at time of school entry was obtained and was treated as a continuous variable.

Sex. A measure of sex was also secured. Males composed 92 percent of the sample and females 8 percent.

First Choice. This dichotomous measure indicates whether or not a student received his first choice of the training courses available.

Military Dependent. This is a dichotomous measure indicating whether a student came from a military family.

Student Source. This measure reflected any one of several sources of input into the Technical Training School: Regular Air Force with nonprior service, Regular Air Force retrainee, Regular Reserves, six-month Reserves, Air National Guard, and other. The majority of students were Regular Air Force and nonprior service (78 percent). This variable was reclassified to allow for the use of three categories (Regular Air Force-Nonprior service, Regular Air Force-Retrainee, and "other") with the first two representing dummy variables.

Education Level. This measure originally appeared as twenty-three different categories ranging from less

than high school to a Ph.D. with three professional degrees. On the average, the students represented in our sample had a high school diploma; very few students ranged beyond one or two years of college. Only 3 percent did not possess a high school diploma or equivalent. This variable was recoded as a continuous variable indicating the number of years of education completed. The new scale ranges from less than high school (eleven years) and advances in one-year increments to reflect up to eighteen years of education.

Instructional Variables

Shift. Shift is a dichotomous variable measuring whether courses meet during the day or on a swing shift (i.e., 0600-1500 or 1500-2400 hours, respectively).

Hours Absent. This is a continuous measure of the total number of hours (to the nearest hour) a student is absent from a training course for any reason. Hours absent ranged as follows: 62 percent had zero hours, 22 percent had between one and five hours, 10 percent had between six and eleven hours, and 6 percent had up to sixty-five hours absent.

Tutoring. This continuous variable measures the time an instructor spent tutoring an individual student. Hours tutored ranged as follows: 41 percent received no

tutoring, 31 percent received one to three hours, 21 percent received four to twelve hours, and 7 percent received up to fifty-one hours of tutoring.

Academic Counselings. This variable represents the number of times an instructor counsels an individual during a training course. Duration of sessions vary from a few minutes to about an hour. The number of academic counseling sessions ranged as follows: 81 percent received no counselings, 17 percent were counseled one to four times, and 2 percent were counseled up to nine times.

Nonacademic Counselings. This variable measures the number of nonacademic counselings (for discipline or administrative reasons) occurring during the technical training program. The range of incidents of nonacademic counselings was as follows: 94 percent received no counselings, 5 percent were counseled one or two times, and 1 percent were counseled up to eight times for nonacademic reasons.

Procedures

The data were compiled at the Technical Training School, Chanute Air Force Base, Illinois, from a survey worksheet consisting of three sections (see Appendix C). The initial information was provided by the trainee upon entering a training course. The second portion of the

worksheet was filled out by school administrative personnel from records or interviews. The final portion was completed upon each student's graduation or removal from training.

Analytic Techniques

Correlation and regression analyses are often used to investigate the strength and nature of linear relationships among interval scaled variables. This study researches such relationships by making use of the Pearson Product-Moment Correlation and stepwise multiple regression analysis.

Pearson Product-Moment Correlation. One of the simplest measures of relationships between two interval-level variables is the zero-order or simple correlation (Pearson *r*). If the predictor variables to be investigated in a study are totally independent of one another, paired correlations between each predictor and the criterion can be used to determine relative sizes of predictor validities. But if intercorrelations exist among predictors, a more sophisticated analytic technique is necessary to assess the unique variance in the criterion explained by each individual predictor.

Means, standard deviations, and intercorrelations (Pearson *r*'s) among all variables were calculated. Missing

data were handled by dropping cases containing any missing data (listwise deletion).

Multiple Regression. Multiple regression is a statistical technique facilitating the analysis of the relationships between a criterion variable and a set of predictor variables. It fits a linear combination of predictor variables to the data using the least squares technique resulting in a prediction equation. The model hypothesized to represent the relationship between the criterion and predictor variables has the form:

$$Y = \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k$$

where, Y is the criterion variable;

$X_1, X_2, \ldots, X_k$ is a set of predictor variables;

$\beta_1, \beta_2, \ldots, \beta_k$ are unknown parameters.

To enable the writers to perform hypothesis tests with multiple regression, the following standard statistical assumptions for the random error terms, ε_i, of the linear model above are made. The random variables ε_i, i=1,2,...n are statistically independent, identically distributed, from a population with zero mean, and normally distributed (McNichols, 1979:4-41).

Stepwise Multiple Regression. The basic Statistical Package for the Social Sciences (SPSS) (Nie, Hull

Jenkins, Steinbrenner, and Bent, 1975:320-327) regression subprogram estimates beta coefficients for variables which are initially specified by the user for inclusion in the prediction model. However, unless other analytic techniques are used to determine which variables should be included, the "best" regression equation for a given number of significant variables cannot be obtained. The SPSS supports sequential model building by means of forward selection, backward selection, and stepwise techniques.

Stepwise multiple regression was ultimately chosen to support the main analysis since it incorporates what the researchers considered to be the most favorable elements of the three options as follows:

1. It is initiated with only a constant term in the equation (as in forward selection).

2. It allows independent variables to be entered one by one on the basis of the partial F-test (as in forward selection) to test for significance, with the size of the F-value determining the entry order (largest first).

3. It considers the variables in the model for deletion (as in backward selection). Stopping rules are included in the process which limit the maximum number of iterations to be performed with a specified critical F-value. The user specifies four parameters in the regression design statement which controls the iterative model building process (McNichols, 1979:4-60).

For the present investigation, the critical F to obtain a 95 percent confidence level would normally be 3.84. However, to preclude the possibility of stopping too soon, the authors set the minimum F-level for inclusion at 2.5. This procedure often results in the identification of significant variables which occasionally enter after nonsignificant ones enter due to multicollinearity effects among variables.

The SPSS regression subprogram yields a variety of useful outputs directly from the computer printout. In addition to the nonstandardized and standardized estimates of the regression coefficients, SPSS automatically outputs F-tests for the overall regression equations, partial F-tests for each coefficient in the model, multiple R^2 and adjusted R^2.

Although standardized coefficients (beta weights) can provide insight into the importance of a variety of predictors on a given criterion, this measure of absolute magnitude does not account for those dependencies between independent variables called multicollinearity.

To account for these effects, and to obtain a more complete measure of the relative importance of the predictor variables for the criterion, the method of comparing incremental contributions to R^2 was used. According to Anderson (1974),

> If a given regression equation contains explanatory variables and R^2 is the squared multiple correlation for the regression, the incremental contribution of the hth explanatory variable is given by $R^2-R_h^2$ where R_h is the squared multiple correlation for a regression on the K-1 variables obtained by deleting the hth variable. Thus, for each of the K variables, the incremental contribution to R^2 is the increase in R^2 that would be obtained if that variable had entered the regression last [Anderson, 1974:12].

Therefore, R^2, as furnished by the SPSS printout, provides an extremely useful indicator of the individual contributions of the predictor variables to explain criterion variance. This statistic was employed to evaluate the statistical significance of each predictor entering the regression model.

Regression analysis was initially performed across all training courses for which data on final course grade were available. The courses are as follows: 42732, 42733, 42734, 42735, 47230, 47231 (A, B, C, and D), 47232, 57130, and 12230. Regression was also performed for those training courses having an $N > 50$. Those courses are as follows: 42732, 42733, 42735, 47231A, 47231B, 47231C, 47231D, 47232, 57130, and 12230. The special vehicle maintenance courses, 47231 (A, B, C, and D) were treated as one course in the analysis since the Air Force uses the same score (General) to assign students to these courses. Also, these courses differ only by a specialty shredout suffix (A, B, C, and D) which identifies the type of vehicle the student will specialize in.

CHAPTER 3

RESULTS

Pearson Product-Moment Correlations

Means, standard deviations, and intercorrelations for all variables (except for noncontinuous variables) over the total sample are provided in Table 1. All correlations were based on an identical sample size (N=1358) after cases having missing data had been dropped.

Final course grade was significantly ($p < .001$) correlated with all five test scores. Correlations ranged from a high of $r=.32$ (general score) to a low of $r=.20$ (administrative score).

Final course grade was also significantly correlated with several biographical items. The two strongest relationships with the course grade criterion were age ($r=.08$, $p < .002$) and educational level ($r=.11$, $p < .001$).

Significant ($p < .001$) negative relationships with the four instructional variables were found. Correlations ranged from a high of $r=-.28$ (academic counselings) to a low of $r=-.14$ (hours absent).

Stepwise Multiple Regression

In each of the regression tables that follows, variables entering the regression equation are listed in

TABLE 1

INTERCORRELATIONS AMONG PREDICTOR AND CRITERION VARIABLES

	Variable	M	SD	1	2	3	4	5	6	7	8	9	10	11	12
1.	Final Course Grade	85.89	8.04		20**	31**	32**	20**	31**	28**	-29**	-14**	11**	-18**	08**
2.	Administrative Score	53.99	19.60			32**	45**	12**	57**	-12**	-04	-05	08**	-06*	03
3.	Electronic Score	54.67	19.71				74**	70**	75**	-14**	-18**	-09**	09**	-05	12**
4.	General Score	55.41	18.40					50**	89**	-16**	-18**	-07**	12**	-03	14**
5.	Mechanical Score	53.02	23.23						47**	-07**	-17**	-05	01	-01	13**
6.	Intelligence Score	55.50	18.19							-15**	-16**	-09**	10**	-05	08**
7.	Tutoring	3.33	5.48								43**	09**	-04	11**	01
8.	Academic Counselings	.34	.92									14**	-07*	19**	-07**
9.	Hours Absent	2.35	5.24										-03	08**	-03
10.	Education Level	12.20	.65											-03	35**
11.	Nonacademic Counselings	.05	.29												04
12.	Age	20.00	3.04												

Note: Decimals have been omitted from correlations.

*$p < .05$; **$p < .01$.

the order by which they entered the model. Only significant predictors are given in the table. "Significance" is defined in this case as a significant increment in $\underline{R}^2$ change with the addition of a given predictor. Standardized regression coefficients (beta weights) are provided in the tables to indicate the relative contribution of each predictor variable for predicting the criterion. Nonstandardized regression coefficients are given in Appendix A in a complete regression equation for individual training courses studied and for the entire training sample.

All Courses Combined. Stepwise multiple regression was performed across all training courses. Table 2 provides an overall test for the total sample regressing final course grade against all study predictors. Due to the large sample size associated with the overall test, many predictors entered the regression model at conventional levels of significance (i.e., .05) which possessed little practical significance (in terms of $\underline{R}^2$ change). Strongest relationships ($\underline{p} < .001$) were obtained between final course grade and general test score, academic counselings, tutoring, and nonacademic counseling. These $\underline{R}^2$ changes ranged from .101 to .014. The electronic test score, shift, and hours absent were significant at $\underline{p} < .01$, but with comparatively smaller contributions to the explanation of criterion variance (.008, .006, and .005,

TABLE 2

REGRESSION OVER ALL AFSCs WITH FINAL COURSE GRADE AS THE CRITERION VARIABLE (N=1358)

Step	Variable Entered	Estimated Coefficient (Beta)	R	R^2	R^2 Change
1	General Score	.141	.318	.101	.101***
2	Academic Counselings	-.130	.394	.155	.054***
3	Tutoring	-.154	.421	.177	.022***
4	Nonacademic Counselings	-.110	.437	.191	.014***
5	Electronic Score	.115	.446	.199	.008**
6	Shift	-.081	.453	.205	.006**
7	Hours Absent	-.071	.459	.211	.005*
8	Education Level	.061	.463	.215	.004*
9	Student Source	.055	.467	.218	.003*
10	Administrative Score	.054	.469	.220	.002*

*$p < .05$.

**$p < .01$.

***$p < .001$.

respectively). All other predictors were significant beyond the .05 level of significance. All beta coefficients contained in the table were consistent with the predicted direction of relationships between the predictors and the course grade criterion. More frequent incidences of absenteeism or counselings (both academic and nonacademic) tended to be associated with poorer performance (as reflected in final course grades). However, shift, by entering the model with a negative weight, revealed that day shift students tended to perform significantly worse than swing shift students. Total $\underline{R}^2$ for the entire regression model was .220. This reflects a moderate level of predictive accuracy.

Nondestructive Inspection Specialist Course. Stepwise multiple regression was performed for each training course individually having an N > 50. These supplementary analyses were carried out to isolate those predictors most useful for forecasting success in each unique program. Table 3 provides the test for students (N=80) in the Nondestructive Inspection Specialist course (AFSC 42732) regressing final course grade with all study predictors. Of the five significant ($\underline{p} < .05$) predictors, the electronic score and frequency of tutoring explained the greatest proportion of variance in the criterion (.351 and .103, respectively), with the remaining three predictors

TABLE 3

REGRESSION FOR AFSC 42732 WITH FINAL COURSE GRADE AS THE CRITERION VARIABLE (N=80)

Step	Variable Entered	Estimated Coefficient (Beta)	R	R^2	R^2 Change
1	Electronic Score	.246	.593	.351	.351*
2	Tutoring	-.238	.674	.454	.103*
3	Student Source	.238	.709	.502	.048*
4	General Score	.311	.728	.529	.027*
5	Shift	-.188	.748	.559	.030*

*$\underline{p} < .05$
**$\underline{p} < .01$
***$\underline{p} < .001$

contributing additional increments in $\underline{R}^2$ ranging from .048 to .030. Total $\underline{R}^2$ for the entire model for this group was .559. The negative relationship between the day shift and the criterion was again noted. Also the presence of the student source variable means, according to our coding of this dummy variable, that being a Regular Air Force Retrainee has a positive impact on final course grade. This may imply that airmen who are cross-trained into other career fields perform better either because they are placed in courses more suited to their skills or because of the prior training experience.

Fabrication and Parachute Specialists Course. Table 4 shows the results for students (N=60) in the

TABLE 4

REGRESSION FOR AFSC 42733 WITH FINAL COURSE GRADE AS THE CRITERION VARIABLE (N=60)

Step	Variable Entered	Estimated Coefficient (Beta)	R	R^2	R^2 Change
1	Tutoring	-.400	.477	.227	.227**
2	Electronic Score	.317	.582	.339	.111**
3	Hours Absent	-.216	.619	.384	.045*

*p < .05
**p < .01
***p < .001

Fabrication and Parachute Specialst Course (AFSC 42733). Of the three predictors to enter significantly, frequency of tutoring sessions and electronic aptitude score were the most highly significant (p < .01) and in combination explained over 88 percent of the predictable criterion variance. Tutoring and hours absent again demonstrated a negative impact on final course grade just as in the regression over all AFSCs. From these results, one can expect decrements in performance not only from students who need tutoring sessions, but also for those students having incidents of absenteeism from their respective training courses. The R^2 for the entire model within this group was .384.

Airframe Repair Specialist Course. The results for students (N=296) in the Airframe Repair Specialist Course

(AFSC 42735) are shown in Table 5. The strongest predictors ($p < .01$) were the electronics score and nonacademic counselings. Of the five significant variables, these two alone explained over 80 percent of the relevant criterion variance. The remaining three variables were significant beyond the .05 level and, as a group, predicted an additional 20 percent of the explained criterion variance. In this group the presence of student source, according to the dummy coding of the variable, indicates that Regular Air Force Retrainee again had a positive impact on final course grade. Total R^2 for the complete regression model was .280.

TABLE 5

REGRESSION FOR AFSC 42735 WITH FINAL COURSE GRADE AS THE CRITERION VARIABLE (N=296)

Step	Variable Entered	Estimated Coefficient (Beta)	R	R^2	R^2 Change
1	Electronic Score	.224	.427	.182	.182**
2	Nonacademic Counseling	-.176	.483	.234	.052**
3	Academic Counselings	-.141	.503	.253	.019*
4	General Score	.184	.518	.268	.015*
5	Student Source	.107	.529	.280	.011*

*$p < .05$
**$p < .01$
***$p < .001$

Special Vehicle Maintenance Courses. Table 6 reports regression findings for students (N=84) in the Special Vehicle Maintenance Courses (AFSC 47231 A, B, C, D). These courses train students in the maintenance of fire trucks, refueling vehicles, materials-handling equipment, and towing and service vehicles. These AFSCs were aggregated and treated as one course due to their close similarity as previously explained. The strongest predictor of performance in this analysis was the mechanical aptitude test score. It alone accounted for approximately two-thirds ($\underline{R}^2$=.35; $\underline{p} < .001$) of the total explained criterion variance. This finding indicates that this particular course draws heavily upon mechanical abilities used in the repair and maintenance of motorized vehicles and equipment. Intelligence test score results and frequency of academic counseling explained most of the remaining criterion variance. These predictors were significant beyond the .01 level. Again, shift entered the equation indicating that students on the swing shift performed significantly better than students on the day shift. Total $\underline{R}^2$ for prediction of success in these courses was .533.

General Purpose Vehicle Mechanic Course. For students (N=56) in the General Purpose Vehicle Mechanic Course (AFSC 47232), only the general test score and number of academic counselings were found to be significant

TABLE 6

REGRESSION FOR AFSCs 47231A, 47231B, 47231C, 47231D WITH FINAL COURSE GRADE AS THE CRITERION VARIABLE (N=84)

Step	Variable Entered	Estimated Coefficient (Beta)	R	R^2	R^2 Change
1	Mechanical Score	.361	.592	.350	*350***
2	Intelligence Score	.275	.671	.451	.100**
3	Academic Counselings	-.290	.713	.509	.058**
4	Shift	-.162	.730	.533	.024*

*$p < .05$
**$p < .01$
***$p < .001$

predictors of academic performance ($p < .001$ and $p < .01$, respectively). These findings imply that the best predictor of success for this general mechanic course is that aptitude test score (general score) indicating proficiency for a relatively broad spectrum of general abilities and skills. Total R^2 for the entire model was .432 (see Table 7).

Fire Protection Specialist Course. Table 8 depicts the results for students (N=541) in the Fire Protection Specialist Course (AFSC 57130). The electronic test score was by far the most significant predictor (R^2=.05; $p < .001$) and it explained approximately half of the predictable criterion variance for this group.

TABLE 7

REGRESSION FOR AFSC 47232 WITH FINAL COURSE GRADE AS THE CRITERION VARIABLE (N=56)

Step	Variable Entered	Estimated Coefficient (Beta)	R	R^2	R^2 Change
1	General Score	.443	.533	.284	.284***
2	Academic Counselings	-.394	.657	.432	.147**

*p < .05
**p < .01
***p < .001

TABLE 8

REGRESSION FOR AFSC 57130 WITH FINAL COURSE GRADE AS THE CRITERION VARIABLE (N=541)

Step	Variable Entered	Estimated Coefficient (Beta)	R	R^2	R^2 Change
1	Electronic Score	.189	.225	.050	.050***
2	Academic Counselings	-.110	.278	.077	.027*
3	Shift	-.100	.301	.091	.013*
4	Hours Absent	-.100	.318	.101	.011*
5	Tutoring	-.102	.332	.110	.009*

*p < .05
**p < .01
***p < .001

Four predictors, each having a negative beta weight of similar magnitude, entered at the .05 level of significance. All were instructional variables as termed in this study. Academic counselings and tutoring were negatively related to performance in this course. Also, absences from class and assignment to the day shift were negatively related to performance. These results were consistent with findings for other technical school courses. Total $\underline{R}^2$ for this group was .110 which reflects a moderate level of predictive accuracy.

Aircrew Life Support Specialist Course. Finally, regression was performed on data for students (N=206) in the Aircrew Life Support Specialist Course (AFSC 12230). The strongest predictors of performance were clearly the intelligence score and academic counselings which explained over 90 percent of the total predicted variance in final course grade (.341). These predictors were significant beyond the .001 level (see Table 9). Two other predictors entered this regression model. The presence of student source, according to the dummy coding used, indicated that Regular Air Force Nonprior Service (see Appendix A) was negatively related to success in training (R^2=.018; $\underline{p} < .05$). The model for this type of technical training also showed that education level was positively related ($\underline{R}^2$=.015; $\underline{p} < .05$) to performance.

TABLE 9

REGRESSION FOR AFSC 12230 WITH FINAL COURSE GRADE AS THE CRITERION VARIABLE (N=206)

Step	Variable Entered	Estimated Coefficient (Beta)	R	R^2	R^2 Change
1	Intelligence Score	.377	.453	.206	.206***
2	Academic Counselings	-.307	.556	.309	.103***
3	Student Source	-.132	.571	.327	.018*
4	Education Level	.122	.584	.341	.015*

*$p < .05$
**$p < .01$
***$p < .001$

Evaluation of Hypothesis 1

Findings pertinent to this hypothesis are reported in Tables 2 through 9. This hypothesis predicted that aptitude and IQ tests may be used to predict final course grade.

In the results previously discussed, test scores not only entered in all eight regression models, but also dominated the results by being the most predictive variables in seven of the eight regression analyses. The single instance where a test score did not enter first in a regression model saw an aptitude test (electronics) enter in step 2. In at least half the models, more than

one test score was significant at $p < .05$. Based on these results, Hypothesis 1 is supported.

Evaluation of Hypothesis 2

Results pertaining to this hypothesis are displayed in Tables 2 through 9. This hypothesis stated that biographical items would add significantly to the prediction of final course grade. Biographical variables were present in four of the eight regression models. However, of the six biographical variables in this study, only student source and education level entered into the regression equations. The other biographical items failed to contribute significant predictive validity in the models. Based on these results, Hypothesis 2 is only tentatively supported.

Evaluation of Hypothesis 3

Tables 2 through 9 are again used to evaluate this final hypothesis. It predicted that instructional variables would significantly add to the prediction of final course grade.

In reviewing the regression results relevant to this hypothesis, it has been noted that instructional variables were present in every model reported in this study. The overall model (Table 2) contained all five instructional variables while each of the models developed for the individual courses contained from one to three

instructional variables. These variables explained variance in the criterion beyond that predicted by test scores and biographical data. Based on these findings, Hypothesis 3 is supported.

CHAPTER 4

DISCUSSION, CONCLUSIONS, AND RECOMMENDATIONS

In assessing the results of the stepwise regression models obtained in this study, one general observation concerning the order and pattern of variable types to enter can be made. Test scores generally surpass all other variable types with regard to validity for predicting the criterion. Instructional variables are the next most common variable type to appear in the modeling process, with biographic material tending to enter the prediction model least frequently.

Test Scores

The results reported in Chapter 3 are consistent with the many previous findings that test scores are useful contributors to the prediction of a training success criterion. Of the five ASVAB test scores investigated, the general, mechanical, electronic and intelligence test scores each emerged as the "best predictors" of training success in at least one of the seven training groups studied. But in the regression over all AFSCs, only the general and electronic scores appeared as "best predictors" ($p < .01$). Apparently, these two aptitude test scores

alone measure most of those abilities required for successful completion of these seven technical training school courses.

The general score, while being the most important score in the overall model, was especially indicative of success in training for general vehicle maintenance, airframe repair, and nondestructive inspection specialist duties. These duties probably require a wide, general range of skills. The electronic score, while having secondary importance to the general score in the overall model, consistently entered as a stronger predictor than the general score when both were present in the model for an individual course. The electronic score was especially indicative of success for the airframe repair, fire protection, nondestructive inspection, and parachute specialist course. Therefore, if only one score were available for predictive purposes, the electronic score would be, in the authors' estimation the best single test score for measuring the wide range of skills needed in a majority of these seven training courses. Since the above course titles do not imply that electronic skills would be the greatest asset to a trainee, further investigation into the content of the test that produces the electronic score would be necessary to really determine the reason for its outstanding predictive validity.

The authors compared aptitude test results to the current practices of Air Force Classifiers responsible for the assignment of students to technical training. Due to considerable disagreement between our findings and current Air Force practice, the authors can not unconditionally support the Air Force's continued use of the specific test scores (Appendix B) for the initial assignment of airmen to technical training.

Classification procedures used by the Air Force were in agreement only one out of seven times with the principal findings of this study regarding the predictive validity of various aptitude tests. The one instance of agreement occurred for the Special Vehicle Maintenance Courses (AFSC 47231 A, B, C, D) which were combined and treated in the analysis as one training course. For these courses, the mechanical test score was the best predictor of final course grade.

For the other six training courses, the general, electronic, and intelligence test scores each emerged as the best predictor of training success in at least one of the training groups studied (as was noted earlier). This finding is in stark contrast to Air Force assignment policy which instead employs the mechanical aptitude test (which was the "best predictor" in only one instance). Due to this inconsistency, the Air Force use of scores could not be conclusively supported by this investigation. However,

test scores demonstrating the highest validity for the criterion for each course should be noted. Students having low aptitude scores in a course for which that score is deemed by our study as critical to performance, might receive additional attention early in training to offset the possibility of low performance or eventual attrition from training.

Instructional Variables

The instructional variables, as one might expect, also demonstrated unique contributions to the prediction of final course grade. In attempting to make a general assessment of patterns of importance for these variables, the frequency and order of variable entry into the models were examined. For the seven individual course models, each of the instructional variables entered at least once (see Tables 2 through 9). Academic counselings entered the regressions most often (five times) and had the highest average predictive strength based on typical R^2. Frequency of tutoring entered the models three times as did the dichotomous shift variable. However, frequency of tutoring made greater average contributions to the explanation of criterion variance than did the shift variable. Hours absent entered the regression models twice and made higher average contributions to the explanation of variance than did the nonacademic counselings variable which

entered only once. Based on these findings, one of the best indicators of poor academic performance is the frequency of academic counseling required by a student. Early incidences of academic counselings (even apparently isolated occurrences) may serve as signposts signaling the need for vigilant monitoring of a student's performance through the use of a detailed academic counseling record. Training monitors might also keep close track of the amount of tutoring a student requires, an element which provides an indication of training deficiencies. Since the number of tutoring sessions is related to poor performance, it follows to some extent that tutoring sessions are not having, in all cases, the desired remediation effect upon student performance. Administrators therefore might do well to explore alternatives to traditionally used tutoring techniques.

An especially important and interesting finding surfaced in the analysis with regard to the dichotomous shift variable. Apparently, students on the swing shift (1500-2400 hours) outperform their contemporaries on the day shift. The consistently negative sign of the beta coefficient for the shift variable, when it entered the regression models, supports this interpretation.

Absenteeism from training should also be tracked for its negative effect upon training performance. While military attendance at training is mandatory, even the

"legitimate" reasons for absence may need to come under closer scrutiny. In the case of the number of nonacademic counselings, an instructional variable reflecting possible discipline or adjustment problems, the findings indicated that there is a negative effect on the criterion as well. This measure should also be monitored to forecast training deficiencies.

Biographical Items

Of the six biographical items, only education level and student source demonstrated any significant predictive validity for the final course grade criterion. Education level, in the few instances that it did enter the models, reflected a positive relationship with performance. Its presence in the overall model and one individual course model implied that the more years of education an airman has, the greater the probability of improved training performance. With regard to the impact of student source on training performance, the relationship with the criterion was also clear. In the regression model over all training courses, student source indicated that Regular Air Force Retrainees do better than non-retrainees. In the Non-destructive Inspection Specialist Course, retrainees also exhibited higher performance. And finally, in the Airframe Repair Specialist Course, retrainees once again did better. The authors concluded that student source did

make a significant difference, but that further research would be required as to the specific reasons for the positive impact that these student sources have on the training success criterion.

Conclusions and Recommendations

The primary intent of this research was to investigate the predictive nature of a variety of measures for the achievement of training success. It is therefore worth noting at this time that the above research objective was somewhat limited by the purely geometrical artifact referred to as "restriction of range."

When the sample used in an investigation such as this one has been defined as a selected subgroup of the population of interest, the statistics are most probably not representative of those for the population. Under such conditions certain restrictions must be placed on the generalizability of the findings. In the instance of test score results, because the narrowed population has a different mean and lower standard deviation on the ASVAB test scores than was exhibited for the more general population, there is a limitation that diminishes the probability of significance in the reported results.

Two restrictions resulted from the selection process used by the Air Force. First, direct restriction is the result of selecting on only the general and mechanical scores as reported in Appendix B. The reason for this

difference . that the population has a mean and standard deviation which is different from the values of all enlisted Air Force selectees (Thorndike, 1949:169). This has the effect of decreasing the absolute value of the validity coefficients found in this study (Thorndike, 1949:170). This effect is an artifact of the reduced range (and therefore narrowed standard deviations) of the scores on the tests used by the Air Force to select the population of airmen from which the present sample was drawn. This has the effect of directly reducing the value of any correlations calculated between scores on the test used for selection and criterion scores relative to the value of the correlation between the same two variables in the unrestricted population. It can be seen that this phenomenon must be considered when interpreting our findings on the general lack of significance of the predictors in current use by Air Force Classifiers.

The indirect restriction followed from the interrelation of the general score to the mechanical score, intelligence score, electronic score, and the administrative score. Similar indirect restriction was reflected by the correlation of the mechanical score with the intelligence score, electronic score, and the general score (see Table 1). As a result of this high degree of preselection, the reported validity statistics for the prediction of final course grades are indirectly reduced by some

unknown magnitude. This results in conservative values for the validity coefficients reported for those ASVAB scores which were not used to select the trainees in this sample. Therefore, training monitors or others involved in working with trainees in these courses should keep in mind that our conclusions are conservative. Counselors who use the information from these ASVAB scores should realize that the limitations on our sample probably means that low scores on these measures are better clues for potential academic difficulties than the data themselves would indicate.

ASVAB test scores certainly demonstrated the highest predictive validity for the final course grade and should be used extensively, according to our recommendations to reduce low performance and possible student attrition from training. Managers should note that the ASVAB test scores can be used to predict those students most likely to succeed in training, and the other variables can be used to monitor and evaluate student performance during training. The use of biographical items showed promise for the prediction of training success, but were not as highly predictive as the instructional variables. In the absence of further research on instructional variables of the types used in this study, the predictive validity and practical use of these measures will remain relatively speculative. The authors believe that these items, when combined with test scores, will prove extremely useful

and lead to improved student performance and lower course attrition rates for enlisted military personnel attending Air Force Technical Training.

APPENDICES

APPENDIX A

REGRESSION EQUATIONS

REGRESSION EQUATIONS FOR EACH OF THE REGRESSION MODELS IN TABLES 2-9
(Nonstandardized Regression Weights)

AFSC	Regression Equation
All AFSCs Combined	$Y=79.897+.062X_1-1.135X_2-.226X_3-3.091X_4+.047X_5 -1.696X_6-.109X_7+.746X_8+1.968X_9$
42732	$Y=77.454+.069X_5-.253X_3+4.179X_9+.110X_1-2.541X_6$
42733	$Y=84.826-.582X_3+.098X_5-.436X_7$
42735	$Y=72.293+.105X_5-3.378X_4-1.928X_2+.092X_1+5.127X_9$
47231 (A, B, C, D)	$Y=75.460+.121X_{11}+.112X_{12}-1.338X_2-2.294X_6$
47232	$Y=73.264+.164X_1-2.724X_2$
57130	$Y=86.817+.075X_5-1.008X_2-1.933X_6-.138X_7-.452X_3$
12230	$Y=79.034+.133X_{12}-3.292X_2-2.728X_{13}+.956X_8$

Where:

- Y = Final Course Grade
- X_1 = General Score
- X_2 = Academic Counselings
- X_3 = Tutoring
- X_4 = Nonacademic Counselings
- X_5 = Electronic Score
- X_6 = Shift (Enter 0 for day shift, enter 1 for swings)
- X_7 = Hours Absent
- X_8 = Education Level (11 years enter 1, 12 years enter 2, 13 years enter 3, 14 years enter 4, 15 years enter 5, 16 years enter 6, 17 years enter 7, 18 years enter 8)
- X_9 = Regular A.F. Retrainee (if Regular A.F. Retrainee, enter 1; otherwise, enter 0)
- X_{10} = Administrative Score
- X_{11} = Mechanical Score
- X_{12} = Intelligence Score
- X_{13} = Regular A.F. Nonprior (if Regular A.F. Nonprior, enter 1; otherwise, enter 0)

APPENDIX B

TECHNICAL TRAINING COURSE SELECTOR SCORES

TECHNICAL TRAINING COURSE SELECTOR SCORES

AFSC	Job Title	Selector*
12230	Aircrew Life Support Specialist	G-30
42732	Nondestructive Inspection Specialist	G-35
42733	Fabrication and Parachute Specialist	M-30
42734	Metals Processing Specialist	M-35
42735	Airframe Repair Specialist	M-35
47230	Base Vehicle Equipment Mechanic	M-35
47231A	Special Vehicle Mechanic (Firetrucks)	M-30
47231B	Special Vehicle Mechanic (Refueling Vehicles)	M-30
47231C	Special Vehicle Mechanic (Materials Handling Equipment)	M-30
47231D	Special Vehicle Mechanic (Towing and Servicing Vehicles)	M-30
47232	General Purpose Vehicle Mechanic	M-35
57130	Fire Protection Specialist	G-35

*The selector codes G (general) and M (mechanical) are used by the U.S. Air Force as minimum scores for student assignment into each of the specified training courses. These codes are outlined in Air Force Regulation 39-1.

APPENDIX C

TECHNICAL TRAINING SURVEY WORKSHEET

3340TCHTG SURVEY WORKSHEET

Part 1

INSTRUCTIONS: To be completed by the trainee upon entry into a course.

1. Course entered into: 3ABR _____ (Code 6) ___ ___ 1-2

2. Last 4 digits of your SSAN: __ __ __ __ 4-7

3. Were you a military dependent prior to entering service?
___ Yes - 0 ___ No - 1 9

4. Was this course your first choice for AF training?
___ Yes - 0 ___ No - 1 11

Part 2

INSTRUCTIONS: To be completed from records or by personal reference to the individual if the information is not available.

5. Aptitudes:
Administrative __ __ 13-14
Electronic __ __ 16-17
General __ __ 19-20
Mechanical __ __ 22-23

6. AFQT/AQE: __ __ 25-26

7. Sex: __Female - 0 __Male - 1 28

8. Age: __ __ 30-31

9. Student Source: (Code 1) __ 33

10. Course Mode: Group paced - 0 Self paced - 1 Lock Step - 2 35

11. Shift Attended: S - 0 T - 1 37

12. Educational Attainment: (Code 2) __ __ 39-40

Part 3

INSTRUCTIONS: To be completed following student's graduation or removal from training.

13. SIA Hours total (to nearest): __ __ 42-43

14. Number of Academic counselings: __ __ 45-46

15. Number of Nonacademic counselings: __ __ 48-49

16. Total hours of absence (to nearest): __ __ 51-52

17. Disposition: (Code 3) __ 54

18. Final Course Grade: __ __ __ 56-58

19. Total Course Time: __ __ __ 60-62

3340 TCHTG SURVEY WORKSHEET
Part 4

INSTRUCTIONS: To be completed by the block instructor following student's successful completion or removal from training

	Block Nr.	Block Grade	Block Fails	Total Hrs Required	Instructor's Name	Codes: (4) Instructor Code	(5) Student Life Style	Student Detail
CARD 2	I	9-11	13	15-17		19-22	24	26
	II	28-30	32	34-36		38-41	43	45
	III	47-49	51	53-55		57-60	62	64
CARD 3	IV	9-11	13	15-17		19-22	24	26
	V	28-30	32	34-36		38-41	43	45
	VI	47-49	51	53-55		57-60	62	64
CARD 4	VII	9-11	13	15-17		19-22	24	26
	VIII	28-30	32	34-36		38-41	43	45
	IX	47-49	51	53-55		57-60	62	64

Code 4

On base (not married) - 0
On base (married) - 1
Off base (not married) - 2
Off base (married but alone) - 3
Off base (married with spouse) - 4

Code 5

Student leader - 0
Drill team - 1
Drum and bugle - 2
None - 3

DATA CODING SYSTEM

24 June 1981*

Item 9, Code 1

A1 Regular AF (Nonprior service)	- 1
A2 Regular AF (Retrainne)	- 2
R Regular Reserves	- 3
S 6-Month Reserves	- 4
C Air National Guard	- 5
Other	- 6

Item 12: Code 2

A Less than High School	-01
B Completed GED - no cert.	-02
C GED certified	-03
D High School diploma	-04
E Some post secondary (15-19 Hrs)	-05
F 1 yr Junior, Comm., or Tech	-06
G 1 yr College (30+ less than 60)	-07
H 2 yrs technical school	-08
I Associate Degree	-09
J 2 yrs without Associate degree	-10
K RN Graduate	-11
L 3 yrs college (90+ hours)	-12
M 4 yrs college - no Post Grad	-13
N College grad with less than +15	-14
O BA with more than 15 hours	-15
P AM, MA, MED with less than 30+	-16
Q AM, MA, MED with 30+ hours	-17
R PhD	-18
S 1st Prof Degree: Acc, Arch, Chir, Pod, Dent, MD, Opt, Osteo, Pharm, Vet, Law, Theo	-19
T 2nd Prof Degree	-20
U 3rd Prof Degree	-21
Y None/NA	-22
Z Unknown	-23

Item 17: Code 3

Graduate	- 0
Academic loss	- 1
Administrative loss	- 2
Withdrawal	- 3
Other	- 4

Life Style: Code 4

On base (not married)	- 0
On base (married)	- 1
On base (unmarried with dependents)	- 5
Off base (not married)	- 2
Off base (married but alone)	- 3
Off base (married with family here)	- 4
Off base (unmarried with dependents)	- 6
Other	- 7

Details: Code 5

Student Leader	- 0
Drill Team	- 1
Drum and Bugle	- 2
None	- 3

Course Entered: Code 6

42732	-00
42733	-01
42734	-02
42735	-03
47230	-04
47231A	-05
47231B	-06
47231C	-07
47231D	-08
47232	-09
57130	-10
92230 (12230)	-11

*Replaces all other undated code lists.

BIBLIOGRAPHY

Anastasi, Anne. Psychological Testing. 3d ed. New York: The Macmillan Company, 1968.

Anderson, Kenneth R. "Predictive Modeling of the Academic Performance of USAF Academy Preparatory School Graduates at the USAF Academy." Unpublished master's thesis, AFIT, School of Engineering, Wright-Patterson AFB OH, 1974.

Asher, James J. "The Biographical Item: Can It Be Improved?" Personnel Psychology, Vol. XXV (1972), pp. 251-269.

Caffrey, Bernard, and Tom Klugh. "Effect of Saturday Classes on Grades and Attendance," Psychological Reports, Vol. XXVII (February 1971), pp. 581-582.

Cascio, Wayne F. "Turnover, Biographical Data, and Fair Employment Practice," Journal of Applied Psychology, Vol. LXI (October 1976), pp. 376-380.

Ehrle, Raymond A. "Quantification of Biographical Data for Predicting Vocational Rehabilitation Success," Journal of Applied Psychology, Vol. XXXXVIII (June 1964), pp. 171-174.

Ghiselli, Edwin E., and Clarence W. Brown. Personnel and Industrial Psychology. McGraw-Hill Series in Psychology. 2d ed. New York: McGraw-Hill Book Company, Inc., 1955.

Goffman, J. M. "Predicting Academic Success for Military Paramedical Personnel," Journal of Medical Education, Vol. L (August 1975), pp. 797-804.

Gussett, James C. "Effect of Friday Afternoon Classes on Grades and Attendance," Psychological Reports, Vol. XXXIX (December 1976), pp. 1037-1038.

Jensen, Harold E., Iris H. Massey, and Lonnie D. Valentine, Jr. Armed Services Vocational Aptitude Battery Development (ASVAB Forms 5, 6, and 7). AFHRL-TR-76-87, Personnel Research Division, Lackland Air Force Base TX, December 1976. AD A037522.

Lecznar, William B. Years of Education as a Predictor of Technical Training Success. PRL-TDR-64-2, Air Force Systems Command, Lackland Air Force Base TX, February 1964. AD 437940.

McClelland, James N., and Fen Rhodes. "Prediction of Job Success for Hospital Aides and Orderlies From MMPI Scores and Personal History Data," Journal of Applied Psychology, Vol. LIII (February 1969), pp. 49-54.

McNichols, Charles W. "An Introduction to Applied Multivariate Data Analysis." Course notes. Air Force Institute of Technology, Wright-Patterson Air Force Base OH, 1979.

Mullins, Cecil J., James A. Ehrles, and Malcolm Ree. Weighting of Aptitude Components Based on Differences in Technical School Difficulty. AFHRL-TR-81-19, Manpower and Personnel Division, Brooks Air Force Base TX, July 1981.

Nie, Norman H., Hadlai C. Hull, Jean G. Jenkins, Karin Steinbrenner, and Dale H. Bent. Statistical Package for the Social Sciences. 2d ed. New York: McGraw-Hill Book Company, 1975.

Proceedings: 17th Annual Conference of the Military Testing Association. U.S. Army Enlisted Evaluation Center, Fort Benjamin Harrison IN, September 15-19, 1975.

Ree, Malcolm J. Function Chief, Test Development and Calibration, Air Force Human Resources Laboratory, Brooks Air Force Base TX. Telephone interview. 15 July 1982.

_____, Cecil J. Mullins, John J. Mathews, and Randy H. Massey. Armed Services Vocational Aptitude Battery: Item and Factor Analysis of Forms 8, 9, and 10. Report AFHRL-81-55, Manpower and Personnel Division, Brooks Air Force Base TX, March 1982.

Schuh, Allen J. "The Predictability of Employee Tenure: A Review of the Literature," Personnel Psychology, Vol. XX (1967), pp. 133-152.

Stone, Thomas H., and Gary T. Athelstan. "The SVIB for Women and Demographic Variables in the Prediction of Occupational Tenure," Journal of Applied Psychology, Vol. LIII (February 1969), pp. 408-412.

Thorndike, Robert L. Personnel Selection. New York: John Wiley and Sons, Inc., 1949.

U.S. Department of the Air Force. Personnel Testing Procedures. AFR 35-8. Chapter 1: "Introduction." Washington: Government Printing Office, 15 March 1978.

Valentine, Lonnie D. Prediction and Air Force Technical Training Success from ASVAB and Educational Background. Manpower and Personnel Division, Brooks Air Force Base TX, May 1977. AD A041735.

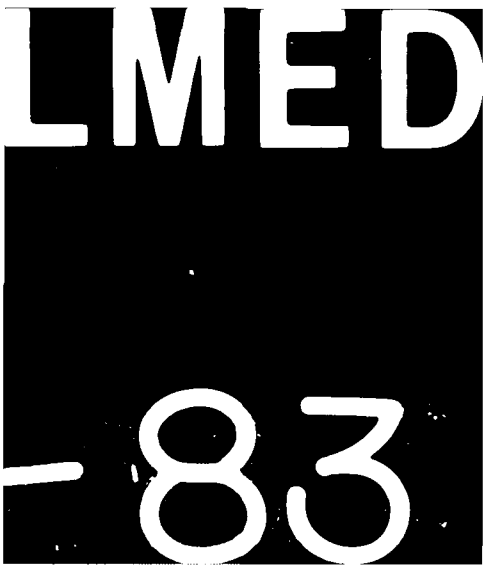
Verna, Steve, and Thomas Mifflin. An Analysis of Marine Corps School Assignment and Performance. CNS 1084. Center for Naval Analysis, Arlington VA, January 1977. AD B021651L.

Vitola, Bart M., and William Alley. Development and Standardization of Air Force Composites for the Armed Services Vocational Aptitude Battery. AFHRL-TR-68-110, Personnel Research Division, Lackland Air Force Base TX, September 1968. AD 688222.

Vitola, Bart M., Cecil J. Mullins, and Paul R. Croll. Validity of Armed Services Vocational Aptitude Battery, Form 1, To Predict Technical School Success. AFHRL-TR-73-7, Personnel Division, Lackland Air Force Base TX, July 1973. AD 767578.

Weeks, Joseph L., Cecil J. Mullins, and Bart M. Vitola. Airman Classification Batteries From 1948 to 1975: A Review and Evaluation. AFHRL-TR-75-78, Personnel Research Division, Lackland Air Force Base TX, December 1975. AD A026470.

Wherry, Robert, and R. Gaylord. "The Concept of Test and Item Reliability in Relation to Factor Pattern," Psychometrika, Vol. VIII (1943), pp. 247-264.


LMED
-83